An algorithm for empirically informed random trajectory generation between two endpoints

Technitis, Georgios ; Weibel, Robert ; Kranstauber, Bart ; Safi, Kamran

Abstract: We present a method to create empirically informed, and thus realistic, random trajectories between two endpoints. The method used relies on empirical distribution functions, which define a dynamic drift expressed in a stepwise joint probability surface. We create random discrete time-step trajectories that connect spatiotemporal points while maintaining a predefined geometry, often based on real observed trajectories. The resulting trajectories have multiple uses, such as to generate null models for hypotheses testing, to serve as a basis for resource selection models, through the integration of spatial context and to quantify space use intensity.

Originally published at:
Technitis, Georgios; Weibel, Robert; Kranstauber, Bart; Safi, Kamran (2016). An algorithm for empirically informed random trajectory generation between two endpoints. In: GIScience 2016: Ninth International Conference on Geographic Information Science, Montreal (Canada), 27 September 2016 - 30 September 2016. s.n., online.

An Algorithm for Empirically Informed Random Trajectory Generation Between Two Endpoints

G. Technitis[1], R. Weibel[1], B. Kranstauber[2,3], K. Safi[2,3]

[1] Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
Email: georgios.technitis@geo.uzh.ch
[2] Department of Migration and Immuno-ecology, Max Planck Institute for Ornithology, Am Obstberg 1, 78315 Radolfzell, Germany
[3] Department of Biology, University of Konstanz, Konstanz, Germany

Abstract

We present a method to create empirically informed, and thus realistic, random trajectories between two endpoints. The method used relies on empirical distribution functions, which define a dynamic drift expressed in a stepwise joint probability surface. We create random discrete time-step trajectories that connect spatiotemporal points while maintaining a predefined geometry, often based on real observed trajectories. The resulting trajectories have multiple uses, such as to generate null models for hypotheses testing, to serve as a basis for resource selection models, through the integration of spatial context and to quantify space use intensity.

1. Introduction

Random trajectories have been increasingly used in movement ecology since their introduction in the early 1980s (Kareiva and Shigesada 1983), gaining significant popularity in the last two decades (Turchin 1998). A wide range of case studies has used the concept, addressing multiple questions related to movement and space use. The majority of the examples found in the literature, however, share one characteristic: the movement has only one restrictive point, the start. Consequently, the simulation is forced to start at a specific location, but can then move according to the set conditions in the given space. In the real world however, this is not always useful: when studying, for example, migration patterns (Codling *et al.* 2010), nest borrowing (Waldeck *et al.* 2008), or fusion of high and low frequency GPS points, the ability to specify an ending point can be crucial. Technitis *et al.* (2015) introduced the Random Trajectory Generator (RTG), an algorithm that enables the user to create randomly varying, possible trajectories between endpoints, based on principles of Time Geography.

In this paper we substantially extend this algorithm. We present a methodology to connect two endpoints by generating empirically informed random trajectories, while preserving the characteristics of the movement. Our approach is based on core theoretical concepts of Time Geography in combination with the Random Walk movement model, and most importantly, we use empirical data to inform our modelling process.

2. Background

Space-time prisms (STP) assist us in calculating the points accessible in space, given the time budget and the maximum speed of an agent (Kuijpers, *et al.* 2010). The calculated path space (in three dimensions defined by x, y and t), and more specifically its 2-D spatial projection, also known as potential path area (PPA), is a homogenous area within which the trajectory lies. The concept of the STP is very intuitive, although it accounts only for the maximum speed of the mover, gives no information regarding the preference of the mover within the given boundaries, and the result is an area, not an individual trajectory.

Bartumeus *et al*. (2005) highlighted the need for movement ecology to add directional persistence into movement modelling, in order to reproduce realistic animal movement, as noted previously by others (Kareiva and Shigesada 1983; Bovet and Benhamou 1988).
Aiming at this gap, Fleming et al. (2014) described a framework that supports the estimation of auto-correlated movement processes, which was later used in home range estimation. Finally, Technitis et al. (2015) presented an algorithm capable of efficiently generating random trajectories between a given origin and destination, with the least bias possible, within the bounds of the STP, honoring speed and time-budget limitations. The significant assumption of this algorithm is that for each step all space-time reachable points are equally probable to be selected. The trajectories derived from this algorithm are all possible, yet not all of them realistic, as they ignore typical movement characteristics of the moving object.
In summary, the direction which movement modeling is taking in ecology seems clear: starting from random walk models, these were successively extended by STP principles and point-to-point constraints. However, what is still missing is the integration of empirically informed movement parameters that can lead to realistic trajectories.

3. Algorithm

Our algorithm generates trajectories with a given set of movement characteristics in stepwise procedure between two consecutive points, in discrete time-steps. The main reasoning for creating probable trajectories in that way is that these consecutive points represent an arbitrary pair of fixes of a trajectory, typically placed at a significant distance due to coarse sampling rate. Each point exerts a different effect on the agent's movement, though of different nature: the probability from the perspective of the fixed starting point (A) expresses the 'local' decisions of a moving individual, such as step-length and direction defining a next position, resulting in a correlated random walk without a priori knowledge about the context of the movement (e.g. resource distributions).

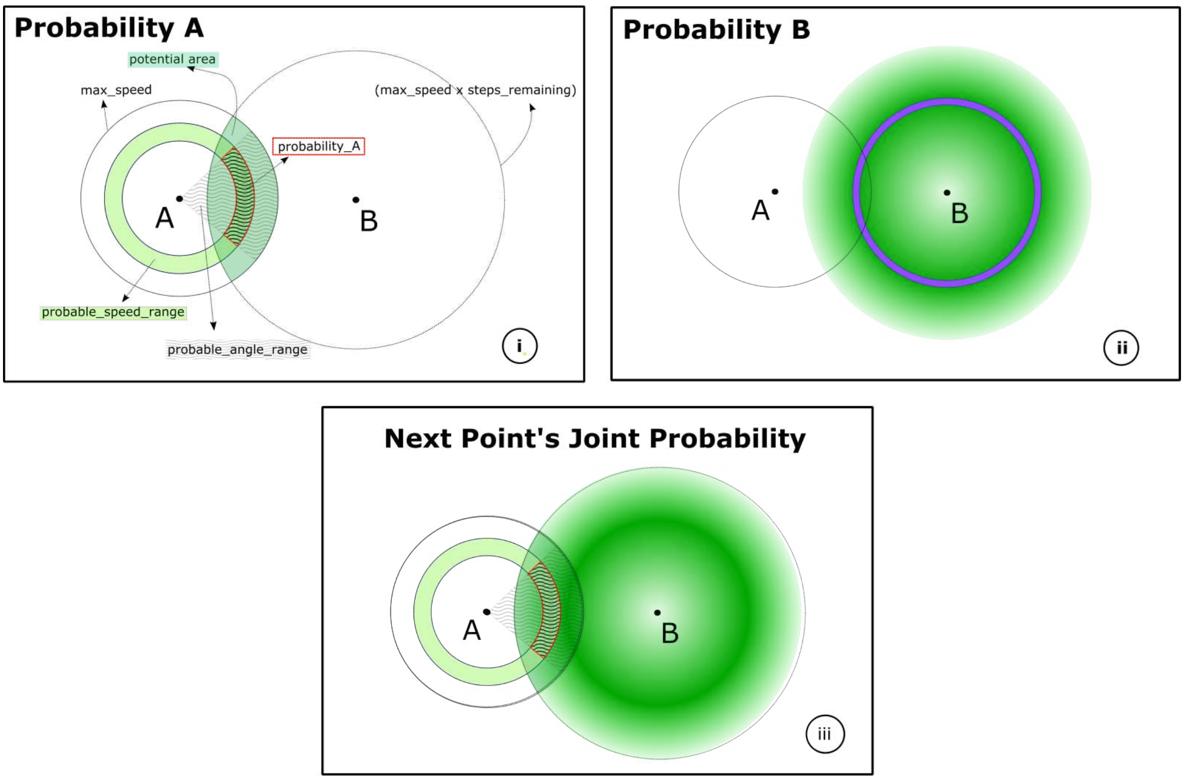

Figure 1. Workflow for calculating the relocation of the mover, (i) shows the effect of the origin (ii) the effect of the destination and (iii) the combined effect.

The probability from the perspective of the ending point (B), represents a gravitational force, forcing the movement towards the desired destination. Depending on the remaining time and the current location, this force is adjusted and acts as a dynamic drift parameter, applying the necessary bias towards the destination. The model can be formally described by a mean-reverting Ornstein-Uhlenbeck process in which “individuals [are] drifting randomly but attracted to an average point” (Smouse et al. 2010). The intensity of the attraction is adjusted over time, ensuring that the mover is within reach of the destination at all times. The flow of the methodology has three steps: firstly, we pre-process the recorded data and we extract the movement parameters of interest. Next, starting from the origin (Figure 1.i), we map the probability of an expected next step based on the value of the turning angle and step lengths. At the same time, we calculate the attraction to the destination for the entire study area (Figure 1.ii) given the number of steps remaining to the end point, resulting in a probability surface for each time step. The number of time-steps (n) is the quotient of the duration of the walk and the user-defined time interval. The last step is to combine the two surfaces into a joint probability (Figure 1.iii) based on which we sample the next relocation point. The procedure is repeated over all time-steps, generating a new random trajectory.

4. Results and Discussion

The proposed algorithm was evaluated on both synthetic and real-life data, with no systematic difference emerging. Here, we report on the results based on three synthetic trajectories representing correlated random walks. We derived the empirical distributions of movement characteristics from these three input trajectories and compared the geometries of the random – simulated - trajectories with the metrics of these input trajectories.

Each input trajectory represented one of the three different behavior, namely opportunistic behavior, foraging and migration. Figure 2 shows the input trajectories in black and the simulated trajectories in grey. We compared the distribution, correlation and persistence of the movement parameters of the random trajectories to the input trajectories to establish that the generated movement patterns match the distributions of the input trajectories.

The resulted trajectories are the first step towards answering questions such as:

- Where was the animal when not observed?
- Do environmental factors affect the movement?
- Which part of the bounded area has the highest occupancy probability, given that the animal was recorded at the points A and B?

The previous version of the algorithm only accounted for the maximum possible speed of the mover, ensuring that it will reach the destination within the given time; the present version respects these conditions and additionally maintains more of the original patterns. The next step will be to optimize the algorithm’s performance and incorporate the effect of the context on the mover’s selections. The beta version of the current work will be released as an open source toolbox under the MOVE-package in R.

Acknowledgements

This research is part of the PhD project of the first author. Funding by the Swiss State Secretariat for Education, Research and Innovation (SERI) through project CASIMO (C09.0167) is gratefully acknowledged.

References

Bartumeus F, da Luz M G E, Viswanathan G M and Catalan J, 2005, “Animal Search Strategies: A Quantitative Random-Walk Analysis.” Ecology 86 (11): 3078–3087.

Bovet P and Benhamou S, 1988, “Spatial Analysis of Animals’ Movements Using a Correlated Random Walk

Model.” *Journal of Theoretical Biology* 131: 419–433. doi:10.1016/S0022-5193(88)80038-9.

Codling E, Bearon R , and Thorn G J, 2010, “Diffusion about the Mean Drift Location in a Biased Random Walk.” *Ecology* 91 (10): 3106–3113. http://.ncbi.nlm.nih.gov/pubmed/21058570.

Fleming C, Calabrese J, Mueller T, Olson K, Leimgruber P and Fagan W F, 2014, “Non-Markovian Maximum Likelihood Estimation of Autocorrelated Movement Processes.” *Methods in Ecology and Evolution* 5 (5): 462–472. doi:10.1111/2041-210X.12176.

Kareiva P M, and Shigesada N, 1983, “Analyzing Insect Movement as a Correlated Random Walk.” *Oecologia* 56 (2-3): 234–238. doi:10.1007/BF00379695.

Kuijpers B, Grimson B and Othman W, 2010, “An Analytic Solution to the Alibi Query in the Space–time Prisms Model for Moving Object Data.” *International Journal of Geographical Information Science* 25 (2): 1–13. doi: 10.1080/13658810902967397.

Smouse P E, Focardi S, Moorcroft P R, Kie J K, Forester J D and Morales J M, 2010, “Stochastic Modelling of Animal Movement.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365 (1550): 2201–2211. doi:10.1098/rstb.2010.0078.

Technitis G, Othman W, Safi K and Weibel R, 2015, “From A to B, Randomly: A Point-to-Point Random Trajectory Generator for Animal Movement.” *International Journal of Geographical Information Science* 8816 (November): 1–24. doi:10.1080/136588100240930.

Turchin P, 1998, *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. *Sinauer*.

Waldeck P, Andersson M, Kilpi M and Öst M, 2008 “Spatial Relatedness and Brood Parasitism in a Female-Philopatric Bird Population.” *Behavioral Ecology* 19 (1): 67–73. doi:10.1093/beheco/arm113.

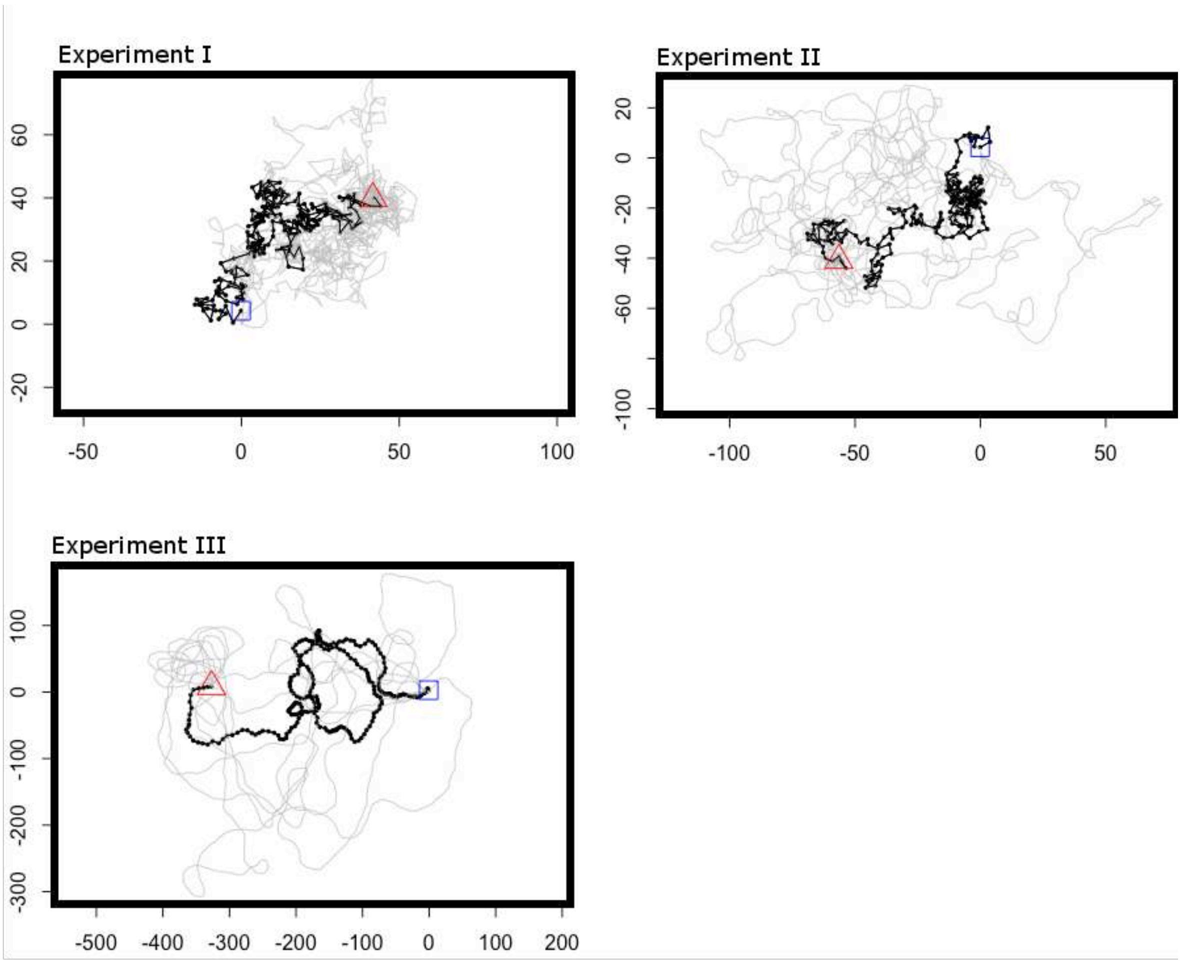

Figure 2. Sample of the simulated trajectories (grey) based on three input trajectories (black). The blue square is the start and the red triangle is the end of the movement. The tree experiments refer to different values of correlation of the movement.